

Planejamento Cirúrgico de Estrabismo Horizontal Utilizando Árvore de Regressão de Múltiplas Saídas

Fernando H. F. Leite¹, João D. S. de Almeida¹, Jorge A. M. Teixeira¹
Luana B. da Cruz¹, Geraldo B. Junior¹, Anselmo Paiva¹

¹Núcleo de Computação Aplicada – Universidade Federal do Maranhão (UFMA)
São Luís – MA – Brasil

{fernandohenrique, jdallyson, jorge, luana.b.cruz, geraldo, paiva}@nca.ufma.br

Abstract. *Strabismus is an eye disease that affects between 1,55% e 3,6% of the population, which can cause irreversible sensory damage to vision. The treatment most severe cases require surgical intervention. Thus, this work uses Regression Trees applied in the form of a single target and multiple targets to indicate the planning of strabismus surgeries. In our most accurate method, Multi-Target Regression Trees, in a single output approach, EAM of 0.49 mm, REQM of 1.49 mm, and was obtained in the indication of the surgical plan for horizontal strabismus.*

Resumo. *O estrabismo é uma patologia oftalmológica que afeta entre 1,55% e 3,6% da população, que pode causar danos sensoriais irreversíveis à visão. O tratamento dos casos mais graves requer intervenção cirúrgica. Desta forma, este trabalho utiliza Árvores de Regressão aplicadas em forma de única saída e múltiplas saídas para indicar o planejamento de cirurgias de estrabismo. Em nosso método mais preciso, Multi-Target Regression Trees, em uma abordagem em única saída, foi obtido o Erro Absoluto Médio de 0,49 milímetros, a Raiz do Erro Quadrático Médio de 1,49 milímetros e R^2 de 0,65 na indicação do plano cirúrgico de estrabismos horizontais.*

1. Introdução

O estrabismo consiste em qualquer desvio do alinhamento binocular e ocorre entre 1,55% e 3,6% da população mundial, a depender de alguns fatores demográficos [Repka et al. 2018]. A classificação do estrabismo se dá pela convergência e divergência dos eixos em relação a um ponto de fixação no sentido horizontal e vertical. No que diz respeito à incidência dessa patologia, ela pode iniciar nos primeiros anos de vida e na idade adulta, e sua origem é variada, podendo resultar de outras doenças.

O estrabismo possui vários métodos de tratamento, que podem ser clínicos ou cirúrgicos e, de acordo com os procedimentos cirúrgicos [Sharma et al. 2017], podem apresentar vários benefícios, entre eles as melhorias psicológicas e funcionais [Ozates et al. 2019]

O procedimento cirúrgico é indicado para a maioria dos casos com desvios acima de 15 Δ (dioptrias prismáticas), devido à presença de sinais e sintomas como diplo-

pia¹, astenopia², torcicolo ocular, nistagmo³, redução de campo visual, comprometimento estético estigmatizante e alterações sensoriais [Astle et al. 2016].

A literatura apresenta poucos trabalhos relacionados a esta pesquisa. Dentre os disponíveis, Almeida et al. (2015) e Silva (2017) propuseram métodos computacionais para planejar automaticamente cirurgia de estrabismo.

No primeiro trabalho, foi utilizado o algoritmo padrão de Máquinas de Vetores de Suporte para Regressão (SVR), treinando um modelo para cada variável-alvo, finalizando a saída da predição pela concatenação de todos os modelos construídos, logo este método tem como objetivo tratar de problemas de única saída, também chamado de *Single-Target* (ST). Utilizaram uma base de dados privada obtida e método de codificação de variáveis categóricas, Label Encoder.

No segundo trabalho, foram desenvolvidos métodos de transformação de problema baseado em SVR para lidar com regressões de múltiplas saídas, a fim de melhor captar inter-relações entre os dados do problema. Foram aplicados os mesmos métodos de normalização e codificação de Almeida (2015), e seus modelos são apresentados como *Multi-Target Regressor Stacking* (MTRS) e *Regressor Chains Corrected* (RCC). Em ambos os métodos ocorrem mais de uma etapa de treinamento, onde no caso do MTRS, o retreinamento do método tem como entrada as características iniciais adicionadas das predições da primeira etapa, e no método RCC os retreinamentos têm como entrada as características iniciais concatenadas de uma predição por vez de cada procedimento cirúrgico entre retreinamentos.

Diferentemente dos trabalhos supracitados, busca-se a aplicação do algoritmo de regressão *DecisionTreeRegressor* [Paper and Paper 2020], baseado no método de *Classification and Regression Trees* (CART) [Ma 2018] e *Multi-Target Regression Trees* (MRTs) [Petković et al. 2017], sendo este adaptado nativamente a uma abordagem de única e múltiplas saídas, assim como a aplicação de diferentes técnicas de codificação para variáveis categóricas como, *Label Encoder* [Li 2018] e *One Hot Encoders* [Potdar et al. 2017]. Com isso, busca-se diminuir o erro das predições das medidas de correção para o procedimento cirúrgico, assim como comparar o comportamento de diferentes técnicas de codificação e métodos para lidar com problemas reais de múltiplas saídas.

As principais contribuições são a aplicação de diferentes métodos de regressão numa abordagem de única e múltipla saída. Assim como a aplicação e comparação de diferentes técnicas de codificação de variáveis categóricas, e finalmente, a diminuição do erro absoluto global médio para um valor inferior a 0,5 mm.

Neste artigo, a Seção 2 apresenta as particularidades do planejamento cirúrgico e do paciente, a Seção 3 descreve as técnicas de regressão. O método proposto e a base de dados com suas características e formas de obtenção estão descritos na Seção 4. As Seções 5 e 6, respectivamente, mostram e discutem os resultados e, por fim, a Seção 7

¹Refere-se aos sintomas de visão dupla, que ocorrem quando uma imagem de um objeto é formada sobre pontos não correspondentes na retina.

²Termo usado para descrever sensações desconfortáveis nos olhos, que geralmente ocorrem quando se trabalha com olhos fixos a uma distância fixa.

³Nistagmo são oscilações repetidas e involuntárias rítmicas de um ou ambos os olhos.

apresenta as conclusões obtidas com o experimento.

2. Planejamento Cirúrgico de Estrabismo

O objetivo da cirurgia do estrabismo é corrigir o desalinhamento dos olhos, isto é, restabelecer a ortotropia-equilíbrio entre os seis pares de músculos extra-oculares e se possível restaurar a Visão Binocular Única (VBU) [Rocha 2016]. Para resolver este problema os procedimentos mais utilizados para correção cirúrgica de acordo com [Rocha 2016] são: o recuo ou retrocesso muscular, miotomia, miectomia, desinserção escleral, alongamento de tendão e sutura de fixação posterior *fadenoperation*, a ressecção ou pregueamento *tucking* do músculo ou do seu tendão e avanço muscular adiante da inserção original. Este trabalho aborda o planejamento dos valores de correção nos procedimentos cirúrgicos de recuo e ressecção no eixo horizontal.

Um fato que vale salientar quanto ao planejamento é que segundo Foschini e Bicas (2001) os planejamentos cirúrgicos produzem resultados coletivos que podem ser satisfatórios, mas os resultados individuais são altamente imprevisíveis. O que confirma a necessidade de desenvolver um método capaz de considerar as inter-relações entre os músculos e suas medidas de correção.

E como particularidades para realizar o planejamento, Khurana (2018) diz que o planejamento do tratamento do estrabismo é um processo complexo e baseia-se em diversos fatores, como idade do paciente, idade de início do estrabismo, condições sensoriais (acuidade visual, presença de ambliopia, supressão retiniana anômala), condições motoras (direção do desvio, variação das medidas de posições, magnitude do ângulo de desvio, concomitância ou incomitância) e etiologia (primários ou secundários à baixa visão, a paresia dos nervos, etc.) do estrabismo, entre outros fatores.

3. Regressão

A análise de regressão tem o objetivo de verificar a existência de uma relação funcional entre uma variável com uma ou mais variáveis, obtendo-se uma equação que explique a variação da variável dependente Y , pela variação dos níveis das variáveis independentes X_1, X_2, \dots, X_p [Haryadi 2017]. Em outras palavras, o principal objetivo da regressão é examinar se as variáveis independentes são bem-sucedidas na previsão do valores-alvos e, quais delas são preditoras significativas deste resultado. Quando temos uma função de relação com mais de uma variável-alvo a ser obtida a partir dos dados em x , temos uma regressão *multi-output* (regressão com múltiplas saídas). [Silva et al. 2017]

3.1. Regressão com Múltiplas Saídas

Regressão de múltiplos alvos preocupa-se com a previsão simultânea de várias variáveis-alvo contínuas com base no mesmo conjunto x de variáveis de entrada, considerando ainda as suas relações internas.

Dado um grupo de dados de treino D com N instâncias contendo um valor determinado para cada variável $X_1, \dots, X_m, Y_1, \dots, Y_d$, i.e., $D = \{(x^1, y^1), \dots, (x^N, y^N)\}$. Cada instância é caracterizada por um vetor de entrada com m variáveis descritivas ou preditivas, $x^l = (x_1^l, \dots, x_j^l, \dots, x_m^l)$ e um vetor de saída com d variáveis-alvo $y^l = (y_1^l, \dots, y_i^l, \dots, y_d^l)$, com $i \in \{1, \dots, d\}$, $j \in \{1, \dots, m\}$, e $l \in \{1, \dots, N\}$. O objetivo é obter, a partir dos dados de treino D , uma função h que seja capaz de gerar, para cada vetor x de características, um vetor y com d variáveis-alvo [Borchani et al. 2015].

Neste trabalho, serão discutidos métodos baseados árvore de decisão com e sem múltiplos valores-alvo, onde este último pode ser classificado como Método de Transformação de Algoritmo.

3.2. Transformada de Algoritmo

Esta abordagem é baseada na ideia de prever simultaneamente todos os alvos usando um único modelo que seja capaz de capturar a interdependência entre eles. De fato, isso tem várias vantagens sobre os métodos de transformação de problemas: é mais fácil interpretar um único modelo com vários destinos do que muitos modelos com um único destino e garante melhor desempenho preditivo, principalmente quando as metas são correlacionadas [Osojnik et al. 2017]. Neste trabalho, serão apresentados os métodos CART e MRTs, métodos estes baseados em árvore binária.

3.2.1. Classification and Regression Trees

CART é um algoritmo baseado em árvore binária de decisão capaz de processar tanto valores contínuos como categóricos [Ma 2018]. Entre as vantagens da CART está a facilidade de construir, a sua interpretabilidade; além da robustez ao lidar com adição de ruído e variáveis auxiliares, assim como com valores ausentes, como perda mínima de informações [De' Ath 2002].

A ideia principal da CART é realizar o particionamento recursivo dos dados em dois subgrupos exclusivos, com o objetivo de determinar geralmente uma variável associada a um critério que permita a melhor divisão e minimize a impureza das subárvores resultantes, até que uma regra ou critério baseado em tempo ou máximo de impureza seja satisfeito [Basti et al. 2015].

Para reduzir os erros de classificação podem ser utilizadas medidas ou critério como Entropia, Coeficiente de Gini, Método dos mínimos quadrados ou *Twoing* [Omurlu et al. 2014]. O crescimento da árvore [James et al. 2013] dá-se da seguinte forma: considere os dados compostos por p entradas e saídas para cada N instâncias que são x_i, y_i para $i = 1, 2, 3, \dots, N$, com $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$. O algoritmo precisa automaticamente decidir uma variável j de particionamento a partir da aplicação de critério de minimização, assim como os pontos de particionamento. Supondo que haja uma base de dados particionada em M regiões R_1, R_2, \dots, R_m , a função da árvore de regressão é definida pela Equação 1:

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m), \quad (1)$$

onde a constante c_m corresponde a uma modelagem da resposta aplicada em cada região M . Adotando como critério de particionamento a minimização da soma dos quadrados $\sum (y_i - f(x_i))^2$, pode-se assumir que o melhor \hat{c}_m , é a média de y_i em cada região R_m conforme mostra a Equação 2:

$$\hat{c}_m = \text{avg}(y_i \mid x_i \in R_m), \quad (2)$$

Para achar o melhor ponto s de partição e a variável j de particionamento em termos de minimização da soma dos quadrados, o algoritmo procede com uma abordagem *greedy* à

procura pela variável j e pelos pontos s que resolvem a seguinte Equação 3:

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right], \quad (3)$$

onde R_1 e R_2 são soluções das equações que compreende o plano do problema $R_1(j, s) = \{X \mid X_j \leq s\}$ e $R_2(j, s) = \{X \mid X_j > s\}$. E, para cada valor j e s a minimização é resolvida pela Equação 4 e 5:

$$\hat{c}_1 = \text{avg}(y_i \mid x_i \in R_1(j, s)) \quad (4)$$

e,

$$\hat{c}_2 = \text{avg}(y_i \mid x_i \in R_2(j, s)). \quad (5)$$

Após obter o melhor particionamento, a divisão é realizada resultando em duas regiões, esse processo é realizado repetidamente em cada uma das regiões resultantes.

3.2.2. Multi-Target Regression Trees

As árvores de regressão com várias saídas são capazes de prever várias saídas y de uma só vez, tendo como principais vantagens seu tamanho menor que as árvores convencionais, e o fato de identificar melhor as interdependências entre as variáveis y para ser previsto [Breskvar et al. 2018].

MRTs são construídas seguindo as mesmas etapas que CART, ou seja, tendo como única diferença a redefinição da medida de impureza de um nó como a soma do erro ao quadrado sobre a resposta multivariada [De' Ath 2002]. Conforme a Equação 6:

$$\sum_{l=1}^N \sum_{i=1}^d (y_i^l - \bar{y}_i)^2 \quad (6)$$

onde $y_i^{(l)}$ indica o valor da variável de saída y_i para a instância l e \bar{y}_i indica a média de y_i no nó. Cada divisão é selecionada para minimizar a soma do erro quadrático e cada folha da árvore pode ser caracterizada pela média multivariada de suas instâncias, o número de instâncias na folha e seus valores característicos definidores.

4. Materiais e Métodos

Esta seção apresenta informações de origem, conteúdo e características da base de dados utilizada nos treinos. Descreve ainda os métodos de pré-processamento, os algoritmos e parâmetros utilizados na construção e aplicação dos modelos de única e múltiplas saídas na previsão automatizada dos valores de correção, referentes ao plano cirúrgico de estrabismo, fazendo por último um contraponto entre técnicas de pré-processamento.

4.1. Base de Dados

A base de dados usada é proveniente do estudo realizado em [de Almeida et al. 2015] e [Silva et al. 2017], obtida em colaboração com uma clínica médica especializada em

oftalmologia em São Luís do Maranhão, por meio da análise manual de prontuários de pacientes obedecendo a um protocolo de admissão estabelecidos por um oftalmologista especialista em estrabismo. A pesquisa possui autorização do comitê de ética em pesquisa do Hospital Universitário Federal do Maranhão através do parecer n.º 331/09.

O banco de dados consiste em dados extraídos de prontuários médicos de 114 pacientes, a partir dos quais são gerados os modelos de regressão responsáveis pela estimativa dos valores de correção do plano cirúrgico, que são posteriormente testados para avaliação de erros.

Os dados fornecidos pelos especialistas para o planejamento cirúrgico consiste em idade, medidas de desvios na posição primária (PPO) e nas secundárias (INFRA, SUPRA, LEVO e DEXTRO), tipo de desvio, olho fixador, equivalentes esféricos, versões dos músculos oculares Reto Medial (RM), Reto Lateral (RL), Reto Inferior (RI), Reto Superior (RS), Oblíquo Inferior (OI) e Oblíquo Superior (OS) de ambos os olhos, o resultado do exame de fundoscopia, totalizando 23 características.

4.2. Aplicação de Métodos de Regressão

As técnicas utilizadas foram: *Single - Target*, *Multi - Target Regression Stacking*, *Regressor Chains Corrected*, e *Multi - Target Decision Trees*, sendo o último aplicado também em formato de única saída. O *Single - Target* utilizado no trabalho foi aplicado por [de Almeida et al. 2015]. Os modelos *Single - Target*, *Multi - Target Regression Stacking* e *Regressor Chains Corrected* foram aplicados por [Silva et al. 2017].

As técnicas de pré-processamento que foram utilizadas são: a normalização dos valores contínuos no intervalo $[-1, 1]$, para padronizar a distribuição de valores das variáveis, a codificação de variáveis categóricas em valores contínuos pelo processo de *Label Encoder* [Li 2018], que transforma valores categóricos numa sequência de números inteiros distintos, e *One Hot Encoders* [Potdar et al. 2017], que transforma em 0 caso o paciente não possua determinada característica e 1 caso possua.

O treinamento dos modelos de regressão foram realizados sobre pacientes escolhidos de forma aleatória, sendo 90% das instâncias da base para treino e o teste sobre o 10% restantes, utilizando a validação cruzada com *v-fold* igual a 10. Valor comumente utilizado por proporcionar estimativas mais confiáveis de acordo com [Witten and Frank 2002].

O treino dos modelos foi realizado na linguagem *Python* na versão 3.7.7 de sua biblioteca padrão, utilizando o pacote *Scikit-learn* 0.22.1. O algoritmo utilizado foi *DecisionTreeRegressor* utilizando o método dos mínimos quadrados para redução de erros como descrito na Seção 3.2.1.

Nos modelos MRTs, parâmetros específicos de configuração foram estimados usando o método de busca exaustiva (*Grid Search*), a fim de procurar exaustivamente a melhor combinação entre todas as possibilidades do espaço de pesquisa definido previamente. Os parâmetros estimados foram: profundidade (*max_depth*), número máximo de folhas (*max_leaf*), número mínimo de amostras por folha (*min_samples_leaf*), e mínimo de amostras para se dividir (*min_samples_split*). Estes parâmetros têm como objetivo permitir que o regressor possa estar ajustado aos dados de treinamento, de forma a permitir aos modelos estimar valores de correção, com precisão, para novos pacientes.

Tabela 1. Valores *max_depth*, *max_leaf*, *min_samples_leaf*, *min_samples_split* para o método MRTs em única e múltipla saída com *Label Encoder* e *One Hot Encoder*.

	Label Encoder									One Hot Encoder								
	Única Saída								Múltiplas Saídas	Única Saída				Múltiplas Saídas				
	RcRM		RsRM		RcRL		RsRL		RcRM/RsRM/RcRL/RsRL	RcRM		RsRM		RcRL		RsRL		RcRM/RsRM/RcRL/RsRL
	D	E	D	E	D	E	D	E	D/E	D	E	D	E	D	E	D	E	D/E
<i>max_depth</i>	5	3	7	5	6	5	7	1	5	5	5	4	7	73	3	4	1	8
<i>max_leaf</i>	9	5	10	11	11	11	8	2	17	12	11	6	16	9	5	6	2	18
<i>min_samples_leaf</i>	3	7	5	1	1	1	7	8	4	1	1	5	1	1	8	6	1	1
<i>min_samples_split</i>	2	2	2	4	2	3	2	2	2	5	9	2	2	10	2	2	2	8

As variáveis alvos Y neste trabalho são referentes aos procedimentos de Recuo no músculo Reto Medial (RcRM) e Ressecção no músculo Reto Medial (RsRM), assim como Recuo no músculo Reto Lateral (RcRL) e Ressecção no músculo Reto Lateral (RsRL), associados tanto ao olho direito (D) e esquerdo (E).

Os parâmetros estimados para MRTs com aplicações em única e múltiplas saídas referente a cada procedimento cirúrgico tanto em *Label Encoder* como em *One Hot Encoder* estão descritos na Tabela 1, respectivamente.

5. Resultados

Esta seção apresenta os resultados obtidos pela aplicação, dos métodos acima descritos, na base de dados. Os modelos gerados a partir do treinamento anterior são aplicados em toda a base utilizando *v-fold* igual a 10 de forma aleatória e intercalada, respeitando a mesma partição para todos os métodos aplicados. A partir disso, os resultados são avaliados em comparação com o plano realizado pelo médico especialista. As seguintes métricas de avaliação foram utilizadas: o Erro Absoluto Médio (EAM), a Raiz do Erro Quadrático Médio (REQM) e o R-quadrado (R^2). O EAM corresponde à média absoluta da distância entre os valores previstos e os valores reais, o que permite uma melhor interpretação do erro na prática; o REQM corresponde ao desvio padrão médio da amostra utilizada, permitindo entender o quão capaz o regressor é para realizar boas previsões [Silva et al. 2017] e o R^2 corresponde a proporção de variação de uma variável (variável objetiva) explicada por outras variáveis (variáveis explicativas) em regressão, o que permite a interpretação do quão bom o modelo é ao prever dados a partir das variáveis predictoras [Kasuya 2019].

Após a geração e o treinamento de cada modelo, foram realizados testes, obtendo-se 8 valores referentes às medidas de correção em milímetros para cada músculo a que se refere. A partir disso, foram calculados os valores relacionados as métricas EAM, REQM e R^2 , quando comparados com as medidas sugeridas pelo especialista. A Tabela 2 mostra as métricas EAM e REQM e R^2 para cada procedimento cirúrgico e as suas médias globais para os métodos ST, MTRS e RCC, e MRTs usando codificação *Label Encoder*.

A Tabela 3, corresponde aos métodos MRTs aplicados em única e múltipla saída utilizando *One Hot Encoder*, e mostram as métricas EAM, REQM e R^2 para cada procedimento cirúrgico e suas médias globais.

Observa-se que o modelo ST baseados em SVR, utilizando *Label Encoder*, obteve sua média global de EAM abaixo de 1 mm, porém tendo seu pior resultado por procedimento em 1,47 mm. No método MTRS com *Label Encoder* foi possível observar melhora

Tabela 2. EAM, REQM e R^2 para o método ST, MTRS, RCC, MRTs em única e múltiplas saídas, com *Label Encoder*.

	Método ST									
	RcRM		RsRM		RcRL		RsRL		Média	
	D	E	D	E	D	E	D	E		
EAM (mm)	1,12	1,04	0,5	1,32	0,79	0,81	1,47	0,53	0,94	
REQM (mm)	1,38	1,18	0,67	1,56	0,93	1,09	1,48	0,61	1,11	
R^2	0,82	0,83	0,58	0,77	0,76	0,70	-0,18	-0,11	0,52	
	Método MTRS									
	EAM (mm)	0,661	0,660	0,746	0,887	0,886	0,930	0,876	0,755	0,798
	REQM (mm)	1,034	1,078	1,035	1,163	1,445	1,287	1,463	1,570	1,259
R^2	0,86	0,84	0,57	0,83	0,75	0,71	0,25	-0,10	0,59	
	Método RCC									
	EAM (mm)	0,756	0,693	0,905	1,151	0,833	0,822	0,947	0,990	0,887
	REQM (mm)	1,062	1,098	1,176	1,459	1,479	1,186	1,487	1,642	1,324
R^2	0,85	0,83	0,51	0,74	0,64	0,64	0,23	-0,11	0,54	
	Método MRTs Única Saída									
	EAM (mm)	0,345	0,416	0,511	0,400	0,711	0,569	0,837	0,444	0,529
	REQM (mm)	1	1,194	1,652	1,290	1,793	1,579	2,217	1,634	1,545
R^2	0,89	0,85	0,45	0,77	0,59	0,72	0,05	0,22	0,57	
	Método MRTs Múltiplas Saídas									
	EAM (mm)	0,669	0,520	0,478	0,338	0,480	0,662	0,778	0,661	0,573
	REQM (mm)	1,450	1,273	1,276	1,040	1,180	1,521	1,842	1,721	1,413
R^2	0,78	0,82	0,67	0,85	0,82	0,74	0,34	0,13	0,64	

Tabela 3. EAM, REQM e R^2 para o método MRTs em única e múltipla saída com *One Hot Encoder*.

	Método MRTs Única Saída									
	RcRM		RsRM		RcRL		RsRL		Média	
	D	E	D	E	D	E	D	E		
EAM (mm)	0,420	0,432	0,472	0,420	0,680	0,517	0,607	0,396	0,493	
REQM (mm)	1,131	1,081	1,5	1,365	1,848	1,598	1,853	1,544	1,490	
R^2	0,86	0,87	0,55	0,75	0,57	0,71	0,33	0,30	0,62	
	Método MRTs Múltiplas Saídas									
	EAM (mm)	0,631	0,466	0,496	0,2274	0,480	0,590	0,833	0,587	0,545
	REQM (mm)	1,484	1,072	1,303	0,867	1,187	1,478	1,937	1,656	1,373
R^2	0,77	0,87	0,66	0,89	0,82	0,75	0,27	0,20	0,65	

significativa da média global no valor de 0,79 mm, porém, houve uma piora nos procedimentos de RsRM D, RsRL D e RsRL E. Quanto ao modelo RCC, este se posicionou entre os dois anteriores, tendo seus modelos individuais maiores variações. Em relação ao REQM entre os métodos baseados em SVR, o método ST possui os menores valores, logo, apresenta melhor capacidade de prever resultados de forma não dispersa. Já em relação ao R^2 entre os métodos em SVR, foram obtidos através de reprodução do estudo, igualando a mesma forma de treino e teste dos modelos de MRTs, apresentando valores negativos em pelo menos um dos procedimentos, indicando uma não correlação entre os modelos e os dados destes procedimentos. Quantos aos procedimentos em que ocorreu

uma correlação aceitável, o método de MTRS apresentou os melhores resultados entre os métodos baseados em SVR, tendo uma média global aproximada de 59%.

Quanto aos modelos de MRTs, utilizando *Label Encoder*, aplicados em única saída, estes obtiveram na sua média global EAM uma melhora significativa em relação ao modelo ST, com média global de 0,52 mm, estando a maioria de seus modelos individuais abaixo de 0,6 mm. Em relação ao REQM, os mesmos modelos não foram capazes de adaptar-se perfeitamente aos dados, apresentando uma piora quanto sua capacidade de realizar boas previsões em relação aos métodos em SVR. Já em relação ao R^2 , estes não obtiveram valores negativos, embora o procedimento RsRL D tenha como resultado um valor próximo de zero, o que novamente indica uma não correlação entre o modelo e os dados deste procedimento. Quanto aos procedimentos em que ocorreu uma correlação aceitável, estes apresentaram um desempenho um pouco inferior aos de MTRS, porém havendo uma melhora no procedimento de RsRL E que apresenta um R^2 de 0,22 ou 22%.

Nos modelos MRTs utilizando *Label Encoder*, aplicado em múltiplas saídas, foi possível notar uma pequena piora em relação aos modelos de MRTs com *Label Encoder* em única saída, tendo uma média global EAM de 0,57 mm. Observa-se uma melhor homogeneidade das médias individuais, com a maioria dos casos abaixo de 0,7 mm. Em relação ao REQM o método também não foi capaz de se adaptar perfeitamente aos dados, porém apresenta uma leve melhora em relação ao método MRTs com *Label Encoder* em única saída, ainda que inferior aos métodos em SVR. Já em relação ao R^2 , o método não obteve valores negativos, com ampla margem de melhora nos procedimentos de RsRM D, RcRL D e RsRL em ambos os olhos. Na totalidade, entre todos os modelos com *Label Encoder*, este modelo apresentou o melhor desempenho quanto a sua capacidade de se correlacionar com os seus dados preditores, tendo como média global de R^2 o resultado de 0,64 ou 64%.

Para os modelos MRTs utilizando *One Hot Encoder* aplicados em única saída, observa-se a melhor média global EAM entre todos os métodos, com valor de 0,49 mm, com cinco procedimentos abaixo de 0,5 mm e três acima. Seu REQM melhorou em relação ao mesmo modelo com *Label Encode*, mesmo que ainda inferior aos modelos em SVR. Em relação ao R^2 , embora este modelo não tenha a melhor média global desta métrica, apresenta os melhores resultados para os procedimentos de RsRL em ambos os olhos.

Para MRTs aplicados em múltiplas saídas com *One Hot Encoder* novamente uma pequena piora em relação ao MRTs com *One Hot Encoder* em única saída, obtendo média global de 0,54 EAM, tendo seus modelos individuais resultados semelhantes quando não utilizando *One Hot Encoder*, onde a maioria dos casos permanece abaixo de 0,7 mm. Seu REQM apresentou o melhor resultado entre os métodos baseados em MRTs, com média global de 1,373, mesmo que não tenha sido capaz de superar os métodos em SVR. Em relação ao R^2 , este modelo apresenta a melhor média global, com uma leve piora em alguns procedimentos específicos, mas com a melhor homogeneidade de resultados, tendo como média global de R^2 o resultado de 0,65 ou 65%.

Vale pontuar que, embora os métodos MRTs aplicados em múltiplas saídas não possuam a melhor média global, estes ocupam a terceira e quarta melhor previsão. Tendo destaque o modelo com *One Hot Encoder* para o músculo RsRM E, que apresentou a

melhor média individual de todos os experimentos em 0,227 mm de EAM. Além disso, observa-se uma melhora no método MRTs pela codificação em *One Hot Encoder* de cerca de 5,0% na média. Para REQM os métodos em MRTs aplicados em múltiplas saídas conseguiram valores melhores em relação aos métodos equivalentes em uma abordagem de única saída, o que confirma uma melhor capacidade dos modelos em se adaptarem a novas predições. Quanto aos R^2 , os mesmos também foram capazes de se correlacionar melhor, em relação aos seus métodos equivalentes em única saída.

6. Discussão

Os melhores valores de EAM são dados por MRTs em única saída, os melhores R^2 pelos métodos de MRTs em múltiplas saídas, em ambos os casos isso se dá independente da codificação. Já os melhores REQM são obtidos pelos modelos em ST. Isso se dá pela forma que os algoritmos calculam os valores de predição, já que SVR busca encontrar valores num espaço, e penaliza predições fora deste espaço, enquanto a árvore explora a melhor característica que permita a melhor predição em cada subdivisão. Logo, a diferença ocorre devido à dispersão dos dados no treino, que tornou o método ST mais propenso ao erro, principalmente no caso dos procedimentos de RsRL, que são pouco realizados, por conta disso tem em sua maioria como medida de correção o valor zero, incapacitando dessa forma o regressor a realizar boas predições quando as mediadas são diferentes de zero, mesmo que este consigam se adaptar melhor a novas predições. Enquanto isso, os métodos baseados em árvores tendem a se adaptar melhor aos dados do problema garantindo menor erro EAM e traçando melhor correlação com as características preditoras.

Os métodos MTRS apresentaram uma melhora significativa em relação ao método ST, conseguindo extrair alguma significância da 2ª etapa. Isso se dá devido ao método acrescentar valores de predições da primeira etapa, na segunda etapa, o que ajuda a popular o espaço no qual o hiperplano será definido, corrigindo as predições dos modelos ST e por consequência, realizando predições mais precisas.

Quanto à melhora proporcionada pela codificação em *One Hot Encoder*, atribui-se o baixo desempenho em árvore, devido à natureza do algoritmo que possui melhor capacidade para lidar com variáveis categóricas nativamente.

Por fim, embora os erros possam ser pequenos em média para alguns pacientes, tanto o método SVR quanto o método MRTs erram as predições, predizendo valores bem próximos de 0, quando o deveria ser maior, divergindo do especialista. Logo, associa-se a incapacidade do método de prever para alguns pacientes a necessidade de expansão da base de dados, principalmente nos procedimentos cirúrgicos menos realizados, principalmente nos casos de correção cirúrgica dos músculos Retos Laterais, que possuem menor incidência.

7. Conclusão

Este trabalho apresentou a viabilidade da utilização de técnicas de regressão para realização do planejamento de cirurgias de correção de estrabismo, possibilitando prever valores de correção para procedimentos cirúrgicos, a partir da análise de dados médicos.

As contribuições deste trabalho estão na melhora das métricas de EAM e R^2 , proporcionando maior confiabilidade no planejamento automatizado das correções cirúrgicas

de estrabismo, pela aplicação das técnicas de pré-processamento e de algoritmos baseados em árvore de regressão com múltiplas saídas nativas, que buscam captar a interdependência entre suas saídas, demonstrando desta forma menor sensibilidade à dispersão e quantidade de dados. No total, diferentes métodos com abordagem em única e múltiplas saídas for comparados, por onde é possível estimar valores para os procedimentos de ressecção e recuo nos músculos horizontais em ambos os olhos, tendo alguns dos modelos mais sucessos que outros quanto a procedimentos individuais.

Os métodos apresentados tiveram diferenças bem próximas para as médias de EAM e REQM, com diferenças inferiores a 0,3 mm entre os melhores e os piores métodos, com a melhor média global de 0,49 mm. Observa-se ainda uma pequena melhora com a aplicação do pré-processamento de *One Hot Encoder*. A partir disso, nota-se a viabilidade da utilização de regressores baseados em árvore em uma abordagem de única e múltiplas saídas já que ambos apresentam médias bem inferiores ao erro de 1,7 mm que segundo [Schutte et al. 2009] corresponde às médias de erro humano relacionado à precisão cirúrgica.

Para melhorar validação dos resultados visa-se ampliar a base de dados, incluindo dados provenientes de outros especialistas para que permita uma melhor generalização das características. Além disso, pretende-se investigar o planejamento cirúrgico para estrabismo vertical.

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) - Proc. 307210/2018-9.

Referências

- Astle, A., Foulsham, T., and McGraw, P. (2016). The consequences of strabismus and the benefits of adult strabismus surgery. *Optometry in Practice*, 17:121.
- Basti, E., Kuzey, C., and Delen, D. (2015). Analyzing initial public offerings' short-term performance using decision trees and svms. *Decision Support Systems*, 73:15–27.
- Borchani, H., Varando, G., Bielza, C., and Larrañaga, P. (2015). A survey on multi-output regression. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(5):216–233.
- Breskvar, M., Kocev, D., and Džeroski, S. (2018). Ensembles for multi-target regression with random output selections. *Machine Learning*, 107(11):1673–1709.
- de Almeida, J. D. S., Silva, A. C., Teixeira, J. A. M., Paiva, A. C., and Gattass, M. (2015). Surgical planning for horizontal strabismus using support vector regression. *Computers in biology and medicine*, 63:178–186.
- De' Ath, G. (2002). Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology*, 83(4):1105–1117.
- Foschini, R. M. and Bicas, H. E. (2001). Cirurgias de músculos retos horizontais: análise de planejamentos e resultados. *Arquivos Brasileiros de Oftalmologia*, 64(6):523–534.
- Haryadi, S. (2017). The non-intercept linear regression method. *Researchgate*. DOI, 10.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

- Kasuya, E. (2019). On the use of r and r squared in correlation and regression. *Ecological Research*, 34(1):235–236.
- Li, J. (2018). Monthly housing rent forecast based on lightgbm (light gradient boosting) model. *International Journal of Intelligent Information and Management Science*, 7(6).
- Ma, X. (2018). *Using classification and regression trees: A practical primer*. IAP.
- Omurlu, I. K., Ture, M., Unubol, M., Katranci, M., and Guney, E. (2014). Comparing performances of logistic regression, classification & regression trees and artificial neural networks for predicting albuminuria in type 2 diabetes mellitus. *Int J Sci Basic Appl Res*, 16(1):173–87.
- Osojnik, A., Panov, P., and Džeroski, S. (2017). Multi-label classification via multi-target regression on data streams. *Machine Learning*, 106(6):745–770.
- Ozates, S., Ozates, M. E., Can, C. U., Polat, S., Yasar, H. H., Taskale, B., and Gogus, A. K. (2019). Improvement in psychiatric symptoms after strabismus surgery in adolescent patients in long-term follow-up. *British Journal of Ophthalmology*, 103(7):966–970.
- Paper, D. and Paper, D. (2020). Scikit-learn regression tuning. *Hands-on Scikit-Learn for Machine Learning Applications: Data Science Fundamentals with Python*, pages 189–213.
- Petković, M., Džeroski, S., and Kocev, D. (2017). Feature ranking for multi-target regression with tree ensemble methods. In *International Conference on Discovery Science*, pages 171–185. Springer.
- Potdar, K., Pardawala, T. S., and Pai, C. D. (2017). A comparative study of categorical variable encoding techniques for neural network classifiers. *International journal of computer applications*, 175(4):7–9.
- Repka, M. X., Lum, F., and Burugapalli, B. (2018). Strabismus, strabismus surgery, and reoperation rate in the united states: analysis from the iris registry. *Ophthalmology*, 125(10):1646–1653.
- Rocha, M. N. A. M. (2016). Propostas de algoritmos e fórmulas para planejamento cirúrgico dos estrabismos horizontais.
- Rolli Khurana, N. S., Srivastava, R. M., and Agrawal, S. (2018). Planning strabismus surgery. *Strabismus: For every Ophthalmologist*, page 115.
- Schutte, S., Polling, J. R., van der Helm, F. C., and Simonsz, H. J. (2009). Human error in strabismus surgery: quantification with a sensitivity analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 247(3):399.
- Sharma, P., Gaur, N., Phuljhele, S., and Saxena, R. (2017). What's new for us in strabismus? *Indian journal of ophthalmology*, 65(3):184.
- Silva, T. A., de Almeida, J. D. S., Teixeira, J. A. M., and Junior, G. B. (2017). Planejamento cirúrgico de estrabismo horizontal utilizando regressores de múltiplas saídas. In *Anais do XVII Workshop de Informática Médica*. SBC.
- Witten, I. H. and Frank, E. (2002). Data mining: practical machine learning tools and techniques with java implementations. *Acm Sigmod Record*, 31(1):76–77.